Evolution of division of labor in self-organizing spatially structured groups

Irene Bouwman

December 23, 2022

Examiner: Dr. Rutger Hermsen
Second reviewer: Prof. Dr. Rob de Boer
Theoretical Biology and Bioinformatics
Utrecht University

Contents

1 Abstract **3**

2 Layperson's summary **4**

3 Introduction **5**

4 Model **7**

5 Results **9**
5.1 Division of labor between public-good producers and non-producers is selected 9
5.2 The lower the production per producer, the higher the probability to produce 10
5.3 A subset of producers maintains public-good availability in every spatial group 10

6 Discussion **14**

7 Methods **16**
7.1 Model and implementation 16
7.2 Calculation of the reproduction rate 16
7.3 Data collection and analysis 17
7.4 Software and code availability 18

References **19**

8 Supplementary simulations **21**
8.1 Methods 21
8.1.1 Software and code availability 21
8.2 Results 21
8.2.1 Competition within groups 21
8.2.2 Competition among groups 21
8.3 Conclusions 22

9 Supplementary information **23**

1 Abstract

In an evolutionary transition in individuality, a collection of individuals organizes into a group that becomes an individual in its own right. Division of labor among individuals in a group is frequently used as an indicator of individuality on the level of the group. Despite extensive research, the processes and circumstances that contribute to an evolutionary transition in individuality are largely unknown. Notably, many models of evolutionary transitions in individuality take a group-structured population as their starting point. In contrast, we ask whether emergent rather than pre-defined groups can be the onset of an evolutionary transition in individuality. To that end, we use an existing model where individuals evolve a gene for altruism, and thereupon self-organize into spatially structured groups that are units of selection in their own right. An individual that behaves altruistically pays a reproductive cost, but provides a reproductive benefit to others. Here, the model is extended to enable phenotypic differentiation between individuals that display altruistic behavior and individuals that do not. We find that a reproductive division of labor can be selected upon the emergence of spatially distinct groups. Moreover, we show that sterile helpers evolve in the system, and that a small fraction of altruists is sufficient to maintain the emergent group structure. Thus, a pre-defined group structure is not required for selection of division of labor. However, we stress that division of labor is not necessarily selected on the level of the group, but might also increase the fitness of individuals within a group. Future work is needed to determine the roles of within- and among-group selection in the evolution of division of labor. Therefore, we question whether division of labor is a good indicator of group-level individuality.

2 Layperson's summary

Multicellular organisms evolved from single-celled organisms. The different cells in a multicellular organism have different functions, and together they ensure the survival and reproduction of the organism. The transition from autonomous, single-celled organisms to more complex, multicellular organisms is an example of an evolutionary transition in individuality. An evolutionary transition in individuality requires two steps: first, individuals have to organize in groups, and second, the group has to evolve into an individual in its own right. In order to qualify a group as an individual, selection should favor properties that maximize the reproductive success of the group rather than the reproductive success of the individuals in the group. A possible indicator of group-level individuality is a division of labor between individuals that specialize in reproductive tasks and individuals that specialize in helping others at the cost of their own reproduction. Consequently, individuals in the group become dependent on each other for their survival and reproduction.

Despite decades of research, it is not clear what biological processes and environmental circumstances enable an evolutionary transition in individuality. Mathematical modeling is a useful tool to get insight into the processes underlying a biological phenomenon. Remarkably, mathematical models of evolutionary transitions in individuality typically assume that the first step of the transition, group formation, has already occurred. Here, we do not make that assumption. Instead, we use an existing model where individuals that are randomly spread over space self-organize into spatially structured groups. In this study, we ask whether self-organizing groups can be the onset of an evolutionary transition in individuality. To that end, the model is extended to enable the evolution of a division of labor between individuals specialized in reproduction and individuals specialized in helping other individuals. We find that natural selection can indeed favor a division of labor between the individuals in self-organized groups. In particular, we show that individuals can evolve that completely sacrifice their own reproduction to help the other individuals in the group. However, we stress that division of labor may not only improve the reproductive success of the group, but also the reproductive success of individuals within a group. Therefore, division of labor might not necessarily indicate that a group functions as an individual in its own right.

3 Introduction

Biological systems are hierarchically organized. Genes are organized in cells, and cells are organized in multicellular organisms. Multicellular life as we know it today is the result of a major evolutionary transition. Free-living single-celled organisms formed collectives that evolved into complex, multicellular individuals. An evolutionary transition in individuality encompasses two processes: (1) individuals organize into cooperative groups, and (2) cooperative groups become individuals in their own right [1] [2]. Next to the transition from unicellular to multicellular life, the evolution of eusociality in insect species such as honey bees and ants possibly denotes a transition in individuality [3] [4] [5] [6].

Despite decades of research, evolutionary theory does not yet provide an answer to the question what processes and circumstances underlie a transition in individuality [7]. Mathematical models are a useful tool to study how an evolutionary transition in individuality can occur. Notably, many models study transitions of individuality in populations with a pre-defined group structure, thereby implicitly assuming that the first step of the transition - group formation - has already taken place [8] [9] [10] [11] [12] [13]. Alternatively, the formation of groups is imposed by a pre-defined ecological structure [7]. Moreover, models with a pre-defined group structure often allow for variation between but not within groups, so selection can only take place on the level of the group [8] [9] [10] [11] [13]. Notwithstanding the insights obtained from models with a pre-imposed group structure, a full transition in individuality requires the formation of groups and possibly selection among groups. Therefore, a model explaining both formation of cooperative groups and the subsequent evolution of group-level individuality might increase our understanding of evolutionary transitions in individuality.

Few models describe the formation of aggregative groups from a population of diverse, autonomous individuals. However, an individual-based model proposed by Hermsen (2022) shows that group formation can occur spontaneously concurrently with the evolution of altruism [14]. In the model, individuals move, die, and reproduce, and they have the chance to evolve an altruistic trait. Altruistic behavior is costly to the individual that expresses it, but provides a benefit to its interaction partners. If individuals interact with each other locally, but compete with each other more globally, they can evolve altruistic behavior and self-organize in spatially structured groups. This evolution of altruism is a process of evolutionary bootstrapping: existence of spatially distinct groups depends on selection for altruism, but selection for altruism also takes place on the level of the group. Since the emergent groups survive, reproduce via division, and die via extinction, they are units of selection in their own right [14] [15]. Importantly, selection on the level of the emergent groups can favor properties that improve group fitness rather than individual fitness. This might provide an onset for an evolutionary transition in individuality.

In order to qualify a group as an individual organism, selection should favor properties that maximize the reproductive success of the group [16] [17] [18]. Therefore, a reproductive division of labor is often used as an indicator of group-level individuality [19] [20] [21] [2]. For instance, social insect colonies consist of reproductive individuals and helpers that perform non-reproductive tasks [3] [6]. Likewise, multicellular organisms typically have a division of labor between germline cells and somatic cells. Germline cells transfer their genes to the offspring of the organism that they are part of. Somatic cells do not transfer their genes to the next generation, but they are needed for the survival of the multicellular organism. Thus, the survival of individual cells within the multicellular organism depends on the survival of the organism as a whole.

Here, we ask whether self-organizing groups that are units of selection in their own right can be the onset of an evolutionary transition in individuality. Therefore, we aim to model the evolution of a division of labor between the individuals in every self-organizing group. To that end, the model proposed by Hermsen (2022) is extended to enable stochastic gene expression. In the extended model, each individual can stochastically adopt a phenotype in which it displays altruistic behavior, or a phenotype in which it does not. Evolution is allowed to affect (1) the probability that the individual adopts the altruistic phenotype, and (2) the level of its investment in altruism. Division of labor by means of phenotypic differentiation has been observed in multiple bacterial species [22] [23] [24] [25].

Although division of labor between individuals in a group has been modeled before, we here let go of the assumption that a group structure is already present. Instead, we assess whether emergent patterning can give rise to a higher level of individuality. This could enhance insight in the processes and circumstances underlying an evolutionary transition in individuality, including the transition from independent unicellular individuals to complex, multicellular organisms. We show that a reproductive division of labor can be selected in a population with an emergent group structure, and that sterile

helpers can evolve in the population. However, we argue that selection of division of labor does not necessarily act on the level of the self-organized groups. Hence, we question the use of division of labor as an indicator of group-level individuality.

4 Model

This section briefly describes the model used in this study, which is based on the model proposed by Hermsen (2022) [14]. Details are provided in Methods.

We consider a population of individuals in a two-dimensional space. Every timestep, individuals have a probability to (1) move following a diffusion process, (2) die at a fixed rate, or (3) reproduce asexually. Furthermore, individuals can invest in altruistic behavior, which can be interpreted as the production of a public good. Public-good production comes at a reproductive cost for the producer, but provides a reproductive benefit to other individuals. Default parameters of the simulations are summarized in Table 1.

The potential level of public good ϕ that an individual can produce is encoded in its genotype. However, gene expression is stochastic. At birth, every individual has a probability p to adopt the producer phenotype, and alternatively it adopts the non-producer phenotype (Figure 1a). If the individual is a producer, it produces an amount of public good ϕ, and if it is a non-producer it produces nothing.

If we now envision a collection of genotypically identical individuals, a fraction p of that collection will produce an amount of public good ϕ. Hence, the expected level of public good produced per individual can be described as $E = p\phi$. In the simulations presented here, mutations can occur in p and E. The production per producer can be calculated accordingly as $\phi = E/p$. As a result, p purely describes how the production of public good is divided between individuals: in a collection of organisms with given E the value of p does not affect how much public good is produced in total.

Individuals profit from the availability of public good, but producing public good is costly. Moreover, an individual competes with the other individuals in its local neighborhood. Therefore, the reproduction rate of an individual increases with the amount of public good that is available at the position of the individual, but it decreases with the amount of public good that the individual produces and the local population density (Equation (1) in Methods).

Both the level of competition and the amount of public good coming from an individual are modeled as a Kernel Density Estimate (KDE) with a normal distribution as a kernel function. For the competition, the standard deviation σ_{rc} of the kernel function represents the distance up to which individuals can feel significant competition from each other. Hence, σ_{rc} can be interpreted as the scale of competition. For public-good availability, the KDE weighs each individual by the amount of public good that the individual produces. The standard deviation σ_{a} of the normal kernel can be interpreted as the scale of altruism. The more public good an individual produces, the higher the reproductive benefit experienced by the individuals in its neighborhood. However, the reproductive benefit saturates for high levels of public-good availability.

The reproductive cost of public-good production is modeled as a concave function of the amount of public good that an individual produces (Figure 1b). Notably, this introduces an explicit benefit to division of labor. For any population mean level of public-good production E, the mean cost is lower with division of labor ($p < 1$) than without division of labor ($p = 1$).

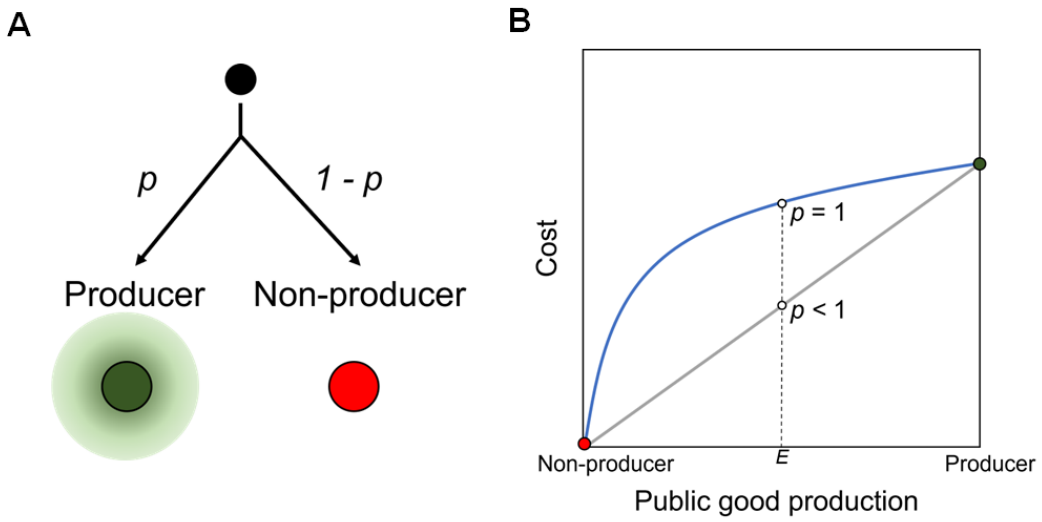

Figure 1: Schematic overview of the phenotypic differentiation model. (A) A newborn individual (black dot) has a probability p to adopt the producer phenotype (green dot), which means that it can increase the public-good availability in its local neighborhood (indicated by the green shade). Alternatively, the individual adopts the non-producer phenotype (red dot). (B) The reproductive cost of public-good production is a concave function of an individual's public-good production (blue line). Therefore, the mean cost of a mean public-good production per individual E (white dots) is lower with a division of labor between producers and non-producers ($p < 1$) than without division of labor ($p = 1$).

Table 1: Default parameters of the simulations.

Symbol	Description	Value
N	Linear size of the grid	1024
d	Death rate	1
g_0	Basal reproduction rate	5
μ_E	Mutation probability E	0.001
m_E	Mean size of mutation in E	0.005
μ_p	Mutation probability p	0.005
m_p	Mean size of mutation in p	0.005
σ_a	Scale of altruism	1
σ_rc	Scale of competition	4
σ_m	Scale of movement	0.08
k_d	Diffusion constant	0.04
b_0	Basal benefit of altruism	1
b_max	Maximum benefit of altruism	5
K	Carrying capacity	40
α	Parameter of the cost function	0.1
κ	Parameter of the cost function	0.1

5 Results

5.1 Division of labor between public-good producers and non-producers is selected

Initially, the population consists of individuals that do not behave altruistically ($E = 0$) and do not have a division of labor ($p = 1$). The initial reproduction rate of an individual is therefore only affected by the local population density. However, individuals that acquire a non-zero expected public-good production due to mutation will likely produce an amount of public good ϕ, and thereby pay a reproductive cost and give a reproductive benefit to their neighbors.

The parameter values have been chosen specifically to enable the evolution of public-good production (Table 1). In this model, evolution of public-good production is facilitated by slow motility and a scale of competition that is larger than the scale of altruism [14]. The results presented here show the outcome of a single representative simulation. (For replicates, see Figure S2.)

In the simulation, the mean level of expected public-good production increases over time and then stabilizes (Figure 2a). As expected, individuals self-organize into spatially distinct groups (Figure 3a, upper panels). Importantly, the evolution of public-good production can be explained by the cumulative effects of selection and drift rather than mutational bias (Figure 2a, red line; see Methods).

While the mean expected public-good production increases over time, the mean probability to adopt the producer phenotype declines as a result of selection (Figure 2b). The mean probability to produce public good stabilizes around 0.1. This indicates that all public good in the system is produced by approximately 10% of the individuals, while the rest of the population consists of non-producers. Indeed, spatial plots of the habitat show an increase in the number of non-producers over time (Figure 3; Video S1). Since the reproductive cost is a concave function of the amount of public good that an individual produces, the mean cost per individual for maintaining a fixed mean level of public-good production decreases with the fraction of producers. Hence, the mean cost paid per individual is lower with a division of labor ($p < 1$) than without a division of labor ($p = 1$). (For mathematical details, see Equation 3 in Methods.) Indeed, the mean cost paid to maintain public-good availability in the stabilized system is approximately 60% lower than it would be without a division of labor (Figure 4). The concave shape of the cost function is required for the evolution of division of labor in this model (Figure S2-S7).

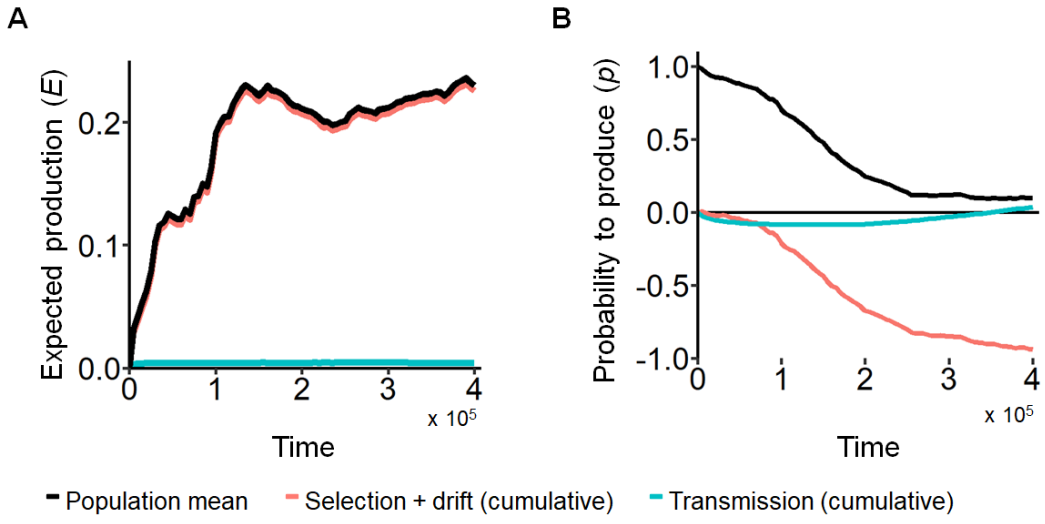

Figure 2: Public-good production and division of labor evolve in the habitat. (A) Population mean expected production and (B) population mean probability to adopt the producer phenotype over time (black), with the relative contributions of selection and drift (red) and transmission (blue).

5.2 The lower the production per producer, the higher the probability to produce

Next, we consider the relationship between the potential level of public good that an individual can produce (ϕ) and the probability that the individual adopts the producer phenotype (p) (Figure 5, snapshot of the simulation after 4×10^5 generations, note the log10 scales). Notably, in Figure 5, the vertical axis represents the amount of public good that an individual would produce if it were to adopt the producer phenotype. The horizontal axis in turn shows the probability that the individual actually adopts the producer phenotype.

The figure shows that if an individual with a low probability to produce nevertheless adopts the producer phenotype, it produces on average more public good than producers with a higher genotypic probability to produce (p). Indeed, the relationship between the production per producer (ϕ) and the probability to produce (p) follows a power law. This indicates that lineages maintain similar expected levels of public-good production (E), but with different fractions of producers (p). In lineages with a low fraction of producers, a single producer produces a large amount of public good, and therefore pays a high reproductive cost.

Interestingly, some lineages generate producers that fully sacrifice their own reproduction rate to produce large amounts of public good (Figure 5, datapoints above the dashed line). That means that the producers from these lineages behave as sterile helpers [3] [26]. As follows from the power law relationship between p and ϕ, the rate at which sterile helpers are produced in a lineage is typically very low: a fraction of less than 0.005 of the individuals in the lineage adopts the producer phenotype. In contrast, in most lineages approximately 1 out of 10 individuals adopts the producer phenotype, and producers are still capable of reproducing themselves, albeit at a lower rate than non-producers.

5.3 A subset of producers maintains public-good availability in every spatial group

So far, we have seen that selection favors a division of labor where a small subset of the population produces high levels of public good. As expected, public-good production leads to the formation of spatially distinct groups. Importantly, the survival of the self-organized groups depends on the availability of public good within each group [14]. Indeed, every distinct spatial group contains producers that maintain public-good availability throughout the group (Figure 6). Notably, groups with more producers do not necessarily have a higher public-good availability than groups with fewer producers. This is in line with the presence of lineages that maintain the same mean production per individual, but with different fractions of producers (Figure 5).

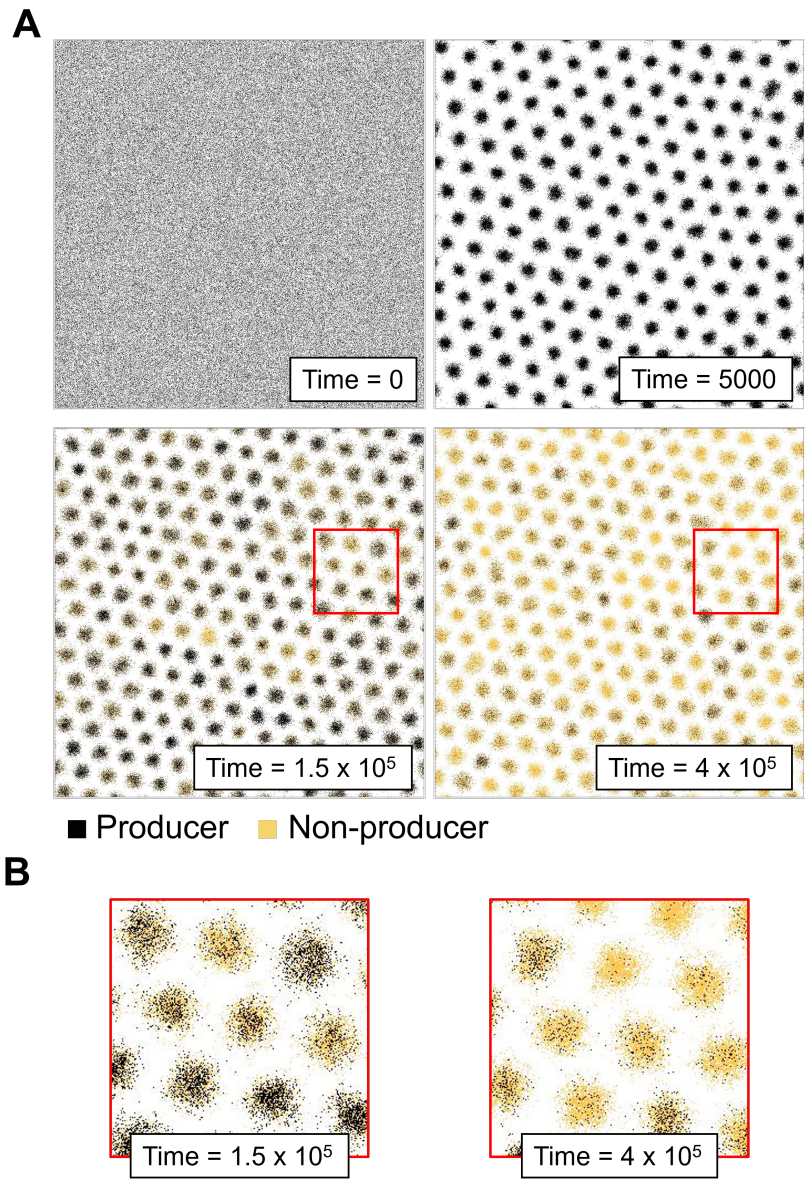

Figure 3: Individuals self-organize into spatially structured groups with producers and non-producers. (A) Spatial plots of the habitat over time. At the start of the simulation, all individuals have the producer phenotype (black, time = 0), but an expected public good production of zero (Figure 2a). Upon the evolution of public good production, individuals self-organize into spatially distinct groups (time = 5000). The number of individuals with the non-producer phenotype (yellow) increases over time, as can be seen in more detail in Video S1. (B) Enlargements of a representative subset of the habitat (indicated by the red squares in A) after 1.5×10^5 (left) and 4×10^5 (right) generations. After 4×10^5 generations, both producers and non-producers are found in every distinct spatial group.

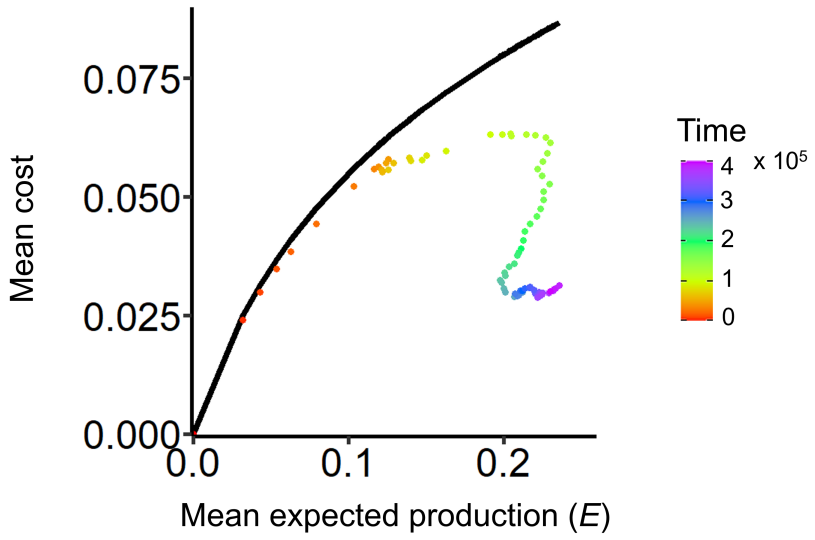

Figure 4: The population mean cost paid to maintain public-good production declines over time. The population mean cost (dots) to maintain a population mean expected public-good production (E) is plotted over time (color scale). The black line shows the population mean cost if all individuals produce the same amount of public good E.

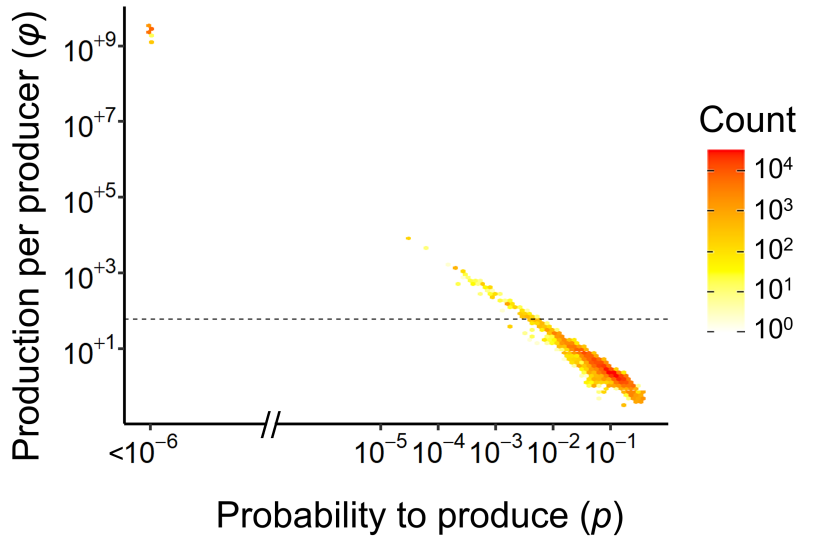

Figure 5: The relationship between the production per producer and the probability to produce follows a power law. The production per producer (ϕ) is plotted versus the probability to produce (p) in the stabilized system (time $= 4 \times 10^5$ in Figure 2). Both producers and non-producers are included in the figure, so the y axis indicates the production if the individual would adopt the producer phenotype. The color scale indicates the number of overlapping individuals. The dashed line indicates the level of production above which producers have a reproduction rate of zero, even if they experience maximum benefit from public-good availability ($\phi = 59$). Note that all scales are logarithmic, and datapoints with $p < 10^{-6}$ are collapsed on the x axis. The data are well described by a power law $\phi \propto p^{-1}$, indicating that despite substantially different values of p individuals tend to have comparable values of E.

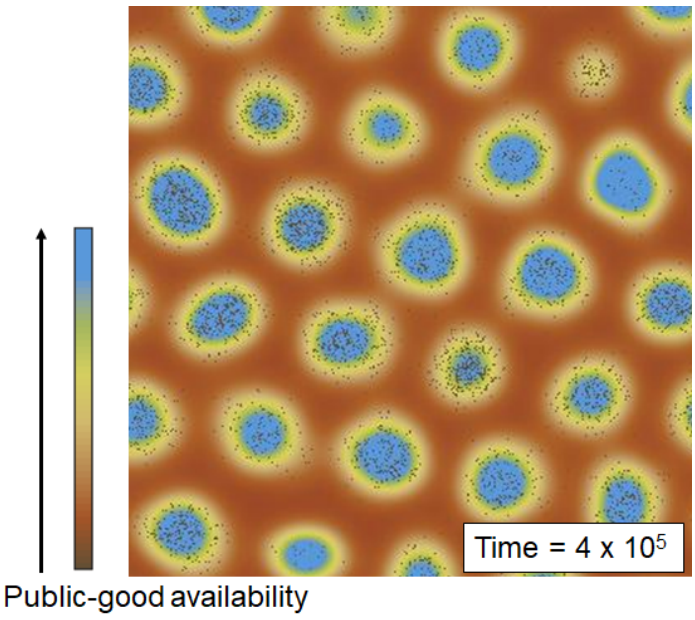

Figure 6: Producers maintain public-good availability throughout each spatial group. Spatial plot of public-good availability in the two-dimensional habitat in the stabilized system (time $= 4 \times 10^5$ in Figure 2). The color scale indicates the level of public good available at every position. Black dots indicate the positions of public-good producers. A representative part of the habitat is shown.

6 Discussion

This study aimed to model the evolution of a division of labor between public-good producers and purely reproductive non-producers in self-organized spatial groups that reproduce in their own right. If the cost of public-good production is a concave function of the amount of public good produced, a division of labor is selected where a small fraction of the individuals produces high levels of public good, while the majority of the population produces nothing. In general, lineages with a low fraction of public-good producers have a higher public-good production per producer than lineages with a higher fraction of producers. Remarkably, sterile helpers evolve that fully sacrifice their own reproduction to invest in public-good production.

In this model, the cost of public-good production is a concave function of the amount of public good that an individual produces. Consequently, the mean cost for maintaining a fixed level of public-good availability is lower when the fraction of individuals that produces public good is lower. However, a group with a low fraction of producers risks stochastic loss of all its producers, especially upon division into two initially smaller groups. The fraction of producers in the stabilized system might be the result of a balance between mean cost reduction and avoiding stochastic loss of public-good availability.

Since producers are free to increase their production at any reproductive cost, different types of division of labor are allowed to evolve in the model. Although non-producers are purely reproductive, producers can either be (1) generalists that invest a part of their reproduction rate in public-good production, or (2) sterile helpers that fully sacrifice their own reproduction rate for public-good production. Interestingly, sterile helpers evolve in the system. This indicates the presence of lineages with a full reproductive division of labor between sterile helpers and pure reproducers. However, most lineages evolve a less strict division of labor where producers can still reproduce themselves, albeit at a lower rate than non-producers. Future work might assess what circumstances affect which type of division of labor evolves.

Selection is measured using the selection differential S of the Price equation (Equation 4 in Methods, [27] [28]). The selection differential S encompasses the effects of both selection and drift. For large population sizes, the effects of drift on the selection differential are expected to be small. In the simulations presented here, the population of groups in the habitat is an order of magnitude 10^3 smaller than the population of individuals. Therefore, the population of groups might be substantially affected by drift. Indeed, large fluctuations in the population means of E, p, and the selection differential S suggest that drift influences the evolution of p and E (Figure 2). Nevertheless, the results presented above are qualitatively reproducible, indicating that random drift does not affect the qualitative outcome of the simulations (Figure S2). Future studies could explore methods to quantify group-level drift and its effects on the evolution of a trait.

Notably, the model considered here does not set an upper limit to the amount of public good that an individual can produce. Thus, lineages that generate sterile helpers can further increase the level of public good produced per producer without paying an additional cost. This is a potential caveat of the model. Future studies can avoid this caveat by limiting individual public-good production to the level where the reproduction rate of the individual becomes zero. For instance, an alternative model could allow individuals to divide the amount of public good that is available to them over investment in public-good production and reproduction (following e.g. [11] and [13]).

Multicellularity has evolved multiple times in the tree of life, both via aggregation of free-living cells and via clonal development of cells that stay connected upon division [29]. The cells in a group that arises through clonal development are genetically identical, whereas cells that form groups via aggregation may have different genotypes. Remarkably, many models on evolutionary transitions in individuality consider only clonal groups [8] [30] [9] [31] [11] [13]. Hence, the potential of non-clonal groups as an onset of an evolutionary transition in individuality has received little attention in modeling studies (but see [26]). In particular, it has often been stated that non-clonal groups are sensitive to disruption by cheaters, and might therefore have limited capacity to undergo an evolutionary transition in individuality [29]. In contrast, we here use a model that demonstrates the emergence of stable non-clonal groups and selection of division of labor, which is in line with recent hypotheses that within-group selection of cheaters might not pose a problem to the evolution of multicellularity [32] [33].

In the model used here, selection can act both on the level of the groups and on the level of the individuals within groups. This is nicely demonstrated by two different mathematical formalizations of the concept of multi-level selection, that are both based on the Price equation [27] [28]: multi-level selection 1 (MLS1) and 2 (MLS2) [34]. In MLS1, the Price equation is applied to the population of

individuals. The selection differential S is decomposed as the sum of selection within groups (S_{within}) and selection among groups (S_{among}). S_{within} is the covariance between the trait value of individuals within a group (e.g. p) and the number of offspring of the individuals. On the other hand, S_{among} quantifies the covariance between the mean trait value of a group and the number of offspring groups. Since the value of p can vary both within and among groups, neither S_{within} nor S_{among} can be assumed to be 0. Therefore, the selection differential of the Price equation cannot be fully attributed to either within- or among-group selection. Alternatively, in MLS2, the Price equation is applied to the population of groups, so the selection differential S equals the covariance between the group mean P and the number of offspring groups W. Hence, the selection differential S quantifies whether groups with a division of labor ($P < 1$) have a higher fitness than groups without a division of labor ($P = 1$). The transmission term T accounts for systematic differences in P between parent and offspring groups, including within-group evolution. Again, since the self-organizing groups are non-clonal, selection for division of labor might indeed not be explained by among-group selection alone [26] [35].

Importantly, division of labor can in this model be beneficial for both individuals and groups. Given a fixed value of E, division of labor reduces the mean cost of a group of individuals. However, the expected value of the cost paid by a single individual with a fixed value of E is also lower when the individual has a lower probability to be a producer. Indeed, a lineage of individuals with a division of labor pays a lower mean cost than a lineage without division of labor, and hence may outcompete other individuals within a group.

Hypothetically, the division of labor in non-clonal groups observed here can be the result of both within- and among-group selection. For instance, within-group selection might drive the emergence of groups with a division of labor, and groups with a division of labor might in turn outcompete groups without a division of labor. To test this hypothesis, the relative contributions of within- and among-group selection to the evolution of division of labor should be measured separately. Calculating among-group selection requires information on the trait value of a group (e.g. the group mean level of p) and the number of offspring groups that a group has in a single timestep. Therefore, this calculation requires a clustering algorithm that identifies self-organized groups, and tracks them over time to calculate their offspring number and trait value. A suitable sufficiently fast clustering algorithm has not yet been implemented for the two-dimensional model used here. This would be an important focus for future work. Nevertheless, we explored experimental setups to study within- and among-group selection in this model (Supplementary simulations, preliminary analyses give non-significant results). Alternatively, separate quantification of within- and among-group selection would be possible if the behavior of the model is studied in a one-dimensional habitat, for which a heuristic clustering algorithm has already been developed [14].

Division of labor is commonly considered to be an indicator, or even a criterion, of group-level individuality [3] [19] [20] [21] [2]. We noted, however, that the evolution of division of labor observed here might not, or only in part, be driven by selection among groups. This suggests that evolution of division of labor can be the result of selection acting on individuals rather than groups. Therefore, division of labor may not be a reliable indicator of group-level individuality for non-clonal groups. Nevertheless, division of labor can increase dependency of group members on each other, which might in turn promote an evolutionary transition in individuality [16] [26]. For example, sterile helpers arise in the model that would not be viable outside the group. In sum, division of labor might be a driver of evolutionary transitions in individuality rather than an indicator or a criterion. Future work is needed to identify traits that indicate group-level individuality. Such traits should be beneficial to the fitness of the group, but not to the fitness of the individuals within the group (e.g. [18] [36]).

Taken together, this study shows selection of division of labor between public-good producers and non-producers in a habitat where individuals self-organize into spatially distinct groups that are subject to natural selection. Thereby, we aimed to integrate group formation and evolution of group-level individuality in one model. In contrast to earlier models, division of labor evolves upon self-organization of groups that reproduce in their own right, rather than in a pre-defined structure of non-reproducing groups [8] [9] [11] [13]. Importantly, we stress that division of labor does not necessarily indicate that a non-clonal group has undergone a transition in individuality. We propose that future studies explore traits that indicate group-level individuality. This can provide further insight into the processes and circumstances contributing to an evolutionary transition in individuality, and enhance our understanding of the origin of complex life.

7 Methods

7.1 Model and implementation

The model is an extended version of the two-dimensional individual-based model presented in Hermsen (2022) [14]. Below is a description of the model and the analyses performed here, focusing on the methodological details of the extensions made in this study.

The model considers a population of individuals that live in a two-dimensional habitat. Individuals can move, reproduce and die, and individuals possess two genetic traits: a probability p to adopt the producer phenotype, and an expected level of public-good production E. Hence, individuals that adopt the producer phenotype produce an amount of public good $\phi = E/p$, whereas non-producers produce nothing. Public-good production is costly, but the availability of public good is beneficial for the reproduction rate of an individual, as described in more detail below.

Before the start of the simulation, individuals are placed randomly in a two-dimensional square grid. The position of every individual can thus be described with an x and a y coordinate. Multiple individuals can be at the same position. A simulation runs for a number of timesteps. During every timestep δt, the state of the system can change. At the start of a timestep, the local density and the availability of public good at every position in the two-dimensional grid are calculated. Then, the fate of every individual i is considered as follows:

1. Death: Every individual has a fixed probability $d\delta t$ to die.
2. Movement: If the individual doesn't die, it can move. Movement is modeled as a diffusion process. For both the x and the y direction, the change in the position of the individual is drawn randomly from a discretized normal distribution with mean 0 and standard deviation $\sqrt{2k_\mathrm{d}\delta t}$. The new x and y coordinates denote the position of the individual in the next timestep, $t + \delta t$
3. Reproduction: The reproduction rate g_i of the individual is calculated using the local density and public-good availability at the position of the individual at the start of the timestep, and the amount of public good that the individual produces (see below). Given that the individual does not die, its probability to reproduce is $g_i\delta t$.
4. Mutation and development of the child: If the individual reproduces, the genetic traits p and E of the offspring mutate with a probability μ_p and μ_E, respectively. The size of the mutation is drawn at random from an exponential distribution with mean m_p for p or m_E for E. The direction of the mutation is positive or negative with equal probability. Next, the offspring adopts the producer phenotype with the probability p that is encoded in its genotype, or it adopts the non-producer phenotype.
5. Position of the child: Initially, the child is placed at the position that was assigned to its parent at the start of the current timestep. However, the child can move relative to this position following the diffusion process described above.

Thus, in every timestep, an individual can (1) die, (2) reproduce, or (3) neither. The three possible outcomes are mutually exclusive. This is a minor difference with the implementation of the model in Hermsen (2022), which is not expected to affect the behavior of the model.

7.2 Calculation of the reproduction rate

The local population density $D(x;\sigma_\mathrm{rc})$ at a position x is calculated as a Kernel Density Estimate (KDE) of a normal distribution with standard deviation σ_rc. Likewise, the public-good availability $A(x;\sigma_\mathrm{a})$ at a position x is a KDE of a normal distribution with standard deviation σ_a, where each individual is weighted by its public-good production. The reproduction rate g_i of an individual i at a position x_i is calculated as:

$$g_i = \max\left(g_0\left(1 - \alpha\phi + \frac{(1-\alpha)\kappa\phi}{\kappa+\phi} + \frac{b_\mathrm{max}A(x_i;\sigma_\mathrm{a})}{b_\mathrm{max}/b_0 + A(x_i;\sigma_\mathrm{a})}\right)\left(1 - \frac{D(x_i;\sigma_\mathrm{rc})}{K}\right), 0\right) \tag{1a}$$

if the individual has the producer phenotype, and

$$g_i = \max\left(g_0\left(1+\frac{b_{\max}A(x_i;\sigma_a)}{b_{\max}/b_0+A(x_i;\sigma_a)}\right)\left(1-\frac{D(x_i;\sigma_{rc})}{K}\right),0\right) \tag{1b}$$

if the individual has the non-producer phenotype.

Every individual has a basal reproduction rate g_0. The reproduction rate of an individual declines linearly with the local population density $D(x_i;\sigma_{rc})$, such that it is zero when the local population density equals K. Naturally, the reproduction rate cannot be smaller than zero. Furthermore, the reproduction rate of an individual decreases with the amount of public good that an individual produces, and increases when public good is available at the position of the individual. The reproductive benefit increases with the amount of public good at the position of the individual $A(x_i;\sigma_a)$ in a linear fashion with slope b_0, but for high levels of public-good availability the benefit saturates at a maximum $b_{\max}$.

The cost term is included in the equation of the reproduction rate only if the individual has the producer phenotype. If $\alpha = 1$, the cost function equals the linear cost function of the original model [14]. However, in the simulations performed here, the reproductive cost of public-good production is modeled as a concave function of the amount of public good that an individual produces ($\alpha < 1$). Initially, an increase $\Delta\phi$ in public-good production results in an increase of $\Delta\phi$ in the cost. However, the cost of an increase in public-good production becomes lower when the individual produces more and more public good. Eventually, the cost increase caused by an increase of $\Delta\phi$ in the public-good production approaches $\alpha\Delta\phi$.

Since the cost of public-good production is a concave function of the amount of public good produced, Jensen's inequality implies that individuals with division of labor ($p < 1$) on average pay a lower cost than individuals without division of labor ($p = 1$) (Figure 7). For a collection of individuals with a given expected public-good production E and probability to produce p, the mean cost C is:

$$C = p\left(\alpha E/p + \frac{(1-\alpha)\kappa E/p}{\kappa + E/p}\right) \tag{2}$$

Hence, for any expected public-good production E, the difference in mean cost C between a collection of individuals without division of labor ($p = 1$) and with division of labor ($p < 1$) is given by:

$$\Delta C = \alpha E + \frac{(1-\alpha)\kappa E}{\kappa + E} - p\left(\alpha E/p + \frac{(1-\alpha)\kappa E/p}{\kappa + E/p}\right) = \frac{(1-\alpha)\kappa E}{\kappa + E} - \frac{(1-\alpha)\kappa E}{\kappa + E/p} \tag{3}$$

It follows from Equation 3 that the smaller the fraction of producers p, the larger the reduction in mean cost compared to a situation without division of labor ($p = 1$).

7.3 Data collection and analysis

Every 5000 generations, the mean population levels of E, p, and the cost are calculated. Results are shown in Figure 2 and 4. The cumulative contribution of selection and drift, and transmission to E and p are calculated using the Price equation [27] [28]:

$$\Delta\overline{x} = S + T \tag{4a}$$

where

$$S = \mathrm{Cov}(x, w) \quad \text{and} \quad T = \overline{w\Delta x} \tag{4b}$$

In equation (4), $\Delta\overline{x}$ denotes the change in the mean value of a trait x during a time interval Δt. The selection differential S equals the covariance between the trait value x and the relative fitness w per individual, and represents the effects of both selection and drift. The transmission bias T is the mean change in the trait value of x between parents and their offspring.

For every cell of in the grid, the density, the summed values of p and ϕ, and the public-good availability $A(x_i;\sigma_a)$ are calculated. These data are used in Figure 3, 5 and 6. For grid cells with 1 individual, the values of p and ϕ of the grid cell reflect the trait values of p and ϕ of the same individual. However, for grid cells with more than 1 individual, the mean p and mean ϕ of the individuals in the grid cell can be calculated from the data, but not the combination of p and ϕ found in each distinct individual. Therefore, only grid cells with 1 individual are included in Figure 5.

The data used to create the figures presented in this report is available at:
https://github.com/ibouwman/altruism/tree/effective-production/Data.

7.4 Software and code availability

The software used to perform the simulations was written in C. The code is available on Github: https://github.com/ibouwman/altruism/tree/effective-production. Figures were made in R version 4.1.1, using the ggplot2 package, and in Wolfram Mathematica.

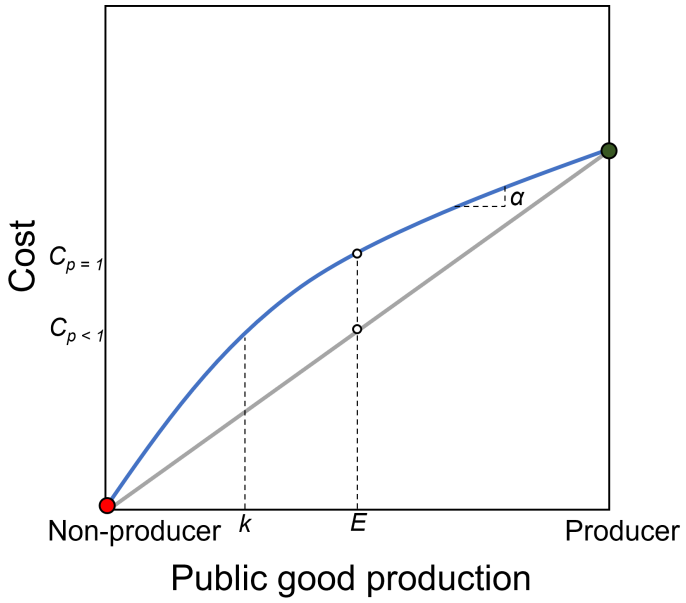

Figure 7: Cartoon of the reproductive cost of public-good production. For any $\alpha < 1$, the cost of public-good production is a concave function of the amount of public good that an individual produces (blue line). Initially, the slope of the cost function is 1, but with increasing levels of public-good production it approaches α. Hence, the mean cost of a mean public-good production per individual E is lower with a division of labor ($C_{p<1}$, on the grey line) than without division of labor ($C_{p=1}$, on the blue line).

References

1. Szathmáry, E. Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences* **112,** 10104–10111. http://www.pnas.org/lookup/doi/10.1073/pnas.1421398112 (2015).
2. West, S. A., Fisher, R. M., Gardner, A. & Kiers, E. T. Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences* **112,** 10112–10119. http://www.pnas.org/lookup/doi/10.1073/pnas.1421402112 (2015).
3. Bernadou, A., Kramer, B. H. & Korb, J. Major Evolutionary Transitions in Social Insects, the Importance of Worker Sterility and Life History Trade-Offs. *Frontiers in Ecology and Evolution* **9.** https://www.frontiersin.org/articles/10.3389/fevo.2021.732907 (2021).
4. Cremer, S. Social immunity in insects. *Current Biology* **29,** 458–463. https://linkinghub.elsevier.com/retrieve/pii/S0960982219303331 (2019).
5. Snir, O. *et al.* The pupal moulting fluid has evolved social functions in ants. *Nature* **612,** 488–494. https://www.nature.com/articles/s41586-022-05480-9 (2022).
6. Taylor, B. A., Reuter, M. & Sumner, S. Patterns of reproductive differentiation and reproductive plasticity in the major evolutionary transition to superorganismality. *Current Opinion in Insect Science* **34,** 40–47. https://linkinghub.elsevier.com/retrieve/pii/S2214574518302050 (2019).
7. Black, A. J., Bourrat, P. & Rainey, P. B. Ecological scaffolding and the evolution of individuality. *Nature Ecology & Evolution* **4,** 426–436. http://www.nature.com/articles/s41559-019-1086-9 (2020).
8. Cooper, G. A., Frost, H., Liu, M. & West, S. A. The evolution of division of labour in structured and unstructured groups. *eLife* **10,** e71968. https://doi.org/10.7554/eLife.71968 (2021).
9. Goldsby, H. J., Dornhaus, A., Kerr, B. & Ofria, C. Task-switching costs promote the evolution of division of labor and shifts in individuality. *Proceedings of the National Academy of Sciences* **109,** 13686–13691. http://www.pnas.org/cgi/doi/10.1073/pnas.1202233109 (2012).
10. Kreider, J. J. *et al.* Resource sharing is sufficient for the emergence of division of labour. *Nature Communications* **13,** 7232. https://www.nature.com/articles/s41467-022-35038-2 (2022).
11. Tverskoi, D. & Gavrilets, S. The evolution of germ-soma specialization under different genetic and environmental effects. *Journal of Theoretical Biology* **534,** 110964. https://linkinghub.elsevier.com/retrieve/pii/S0022519321003842 (2022).
12. Okasha, S. Multilevel Selection and the Major Transitions in Evolution. *Philosophy of Science* **72,** 1013–1025. https://www.journals.uchicago.edu/doi/10.1086/508102 (2005).
13. Yanni, D. *et al.* Topological constraints in early multicellularity favor reproductive division of labor. *eLife* **9,** e54348. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7609046/.
14. Hermsen, R. Emergent multilevel selection in a simple spatial model of the evolution of altruism. *PLOS Computational Biology* **18,** e1010612. https://dx.plos.org/10.1371/journal.pcbi.1010612 (2022).
15. Lewontin, R. C. The Units of Selection. *Annual Review of Ecology and Systematics* **1,** 1–18. https://doi.org/10.1146/annurev.es.01.110170.000245 (1970).
16. Carmel, Y. & Shavit, A. Operationalizing evolutionary transitions in individuality. *Proceedings of the Royal Society B: Biological Sciences* **287,** 20192805. https://royalsocietypublishing.org/doi/10.1098/rspb.2019.2805 (2020).
17. Michod, R. E., Viossat, Y., Solari, C. A., Hurand, M. & Nedelcu, A. M. Life-history evolution and the origin of multicellularity. *Journal of Theoretical Biology* **239,** 257–272. https://linkinghub.elsevier.com/retrieve/pii/S0022519305003875 (2006).
18. Rose, C. J., Hammerschmidt, K. & Rainey, P. B. *Experimental evolution of nascent multicellularity: Recognizing a Darwinian transition in individuality* preprint (Evolutionary Biology, 2020). http://biorxiv.org/lookup/doi/10.1101/2020.03.02.973792.

19. Folse, H. J. & Roughgarden, J. What is an Individual Organism? A Multilevel Selection Perspective. *The Quarterly Review of Biology* **85,** 447–472. https://www.journals.uchicago.edu/doi/full/10.1086/656905 (2010).

20. Michod, R. E. Evolution of individuality during the transition from unicellular to multicellular life. *Proceedings of the National Academy of Sciences* **104,** 8613–8618. https://pnas.org/doi/full/10.1073/pnas.0701489104 (2007).

21. Szathmáry, E. & Maynard Smith, J. The major evolutionary transitions. *Nature* **374,** 227–232. https://www.nature.com/articles/374227a0 (1995).

22. Diard, M. *et al.* Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature* **494,** 353–356. https://www.nature.com/articles/nature11913 (2013).

23. Van Gestel, J., Vlamakis, H. & Kolter, R. Division of Labor in Biofilms: the Ecology of Cell Differentiation. *Microbiology Spectrum* **3,** 3.2.26. https://journals.asm.org/doi/10.1128/microbiolspec.MB-0002-2014 (2015).

24. Hammerschmidt, K., Landan, G., Domingues Kümmel Tria, F., Alcorta, J. & Dagan, T. The Order of Trait Emergence in the Evolution of Cyanobacterial Multicellularity. *Genome Biology and Evolution* **13,** evaa249. https://doi.org/10.1093/gbe/evaa249 (2021).

25. Özkaya, Ö., Xavier, K. B., Dionisio, F. & Balbontín, R. Maintenance of Microbial Cooperation Mediated by Public Goods in Single- and Multiple-Trait Scenarios. *Journal of Bacteriology* **199,** e00297–17. https://journals.asm.org/doi/full/10.1128/JB.00297-17 (2017).

26. Cooper, G. A. & West, S. A. Division of labour and the evolution of extreme specialization. *Nature Ecology & Evolution* **2,** 1161–1167. http://www.nature.com/articles/s41559-018-0564-9 (2018).

27. Price, G. R. Selection and Covariance. *Nature* **227,** 520–521. https://www.nature.com/articles/227520a0 (1970).

28. Price, G. R. Extension of covariance selection mathematics. *Annals of Human Genetics* **35,** 485–490. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1957.tb01874.x (1972).

29. Márquez-Zacarías, P., Conlin, P. L., Tong, K., Pentz, J. T. & Ratcliff, W. C. Why have aggregative multicellular organisms stayed simple? *Current Genetics* **67,** 871–876. https://link.springer.com/10.1007/s00294-021-01193-0 (2021).

30. Gavrilets, S. Rapid Transition towards the Division of Labor via Evolution of Developmental Plasticity. *PLoS Computational Biology* **6,** e1000805. https://dx.plos.org/10.1371/journal.pcbi.1000805 (2010).

31. Maliet, O., Shelton, D. E. & Michod, R. E. A model for the origin of group reproduction during the evolutionary transition to multicellularity. *Biology Letters* **11,** 20150157. https://royalsocietypublishing.org/doi/10.1098/rsbl.2015.0157 (2015).

32. Rainey, P. B. & Kerr, B. Cheats as first propagules: A new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. *BioEssays* **32,** 872–880. https://onlinelibrary.wiley.com/doi/abs/10.1002/bies.201000039 (2010).

33. Veit, W. Evolution of multicellularity: cheating done right. *Biology & Philosophy* **34,** 34. https://doi.org/10.1007/s10539-019-9688-9 (2019).

34. Okasha, S. *Evolution and the Levels of Selection* (Oxford University Press, 2006).

35. Gardner, A. & Grafen, A. Capturing the superorganism: a formal theory of group adaptation. *Journal of Evolutionary Biology* **22,** 659–671. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1420-9101.2008.01681.x (2009).

36. Bourrat, P. Evolutionary Transitions in Individuality by Endogenization of Scaffolded Properties. *The British Journal for the Philosophy of Science.* https://www.journals.uchicago.edu/doi/10.1086/719118 (2022).

8 Supplementary simulations

Quantifying within- and among-group selection separately in a two-dimensional habitat is computationally challenging. A computational method to analyze the relative contributions of within- and among-group selection to the evolution of division of labor in this model is not available. Therefore, experimental setups have been explored to get insight into the roles of within- and among-group selection.

8.1 Methods

In order to calculate the covariance between the value of a trait and fitness on the level of the group, distinct groups must be tracked. However, since automated tracking of self-organized groups is not available, individuals must now be organized in groups before the start of the simulation. In order to mimic the pattern of groups that arises upon self-organization, we imagine that a hexagonal grid with grid constant $a = 8$ is laid over the two-dimensional space before the start of the simulation. Next, all individuals are assigned to one of the grid points at random, and randomly placed within radius $r = 2$ of their grid point. This enables (1) modification of the individuals in a specific group, and (2) tracking group sizes and the traits of the individuals within a specific group. Simulations were performed with the default parameters (Table 1).

Individuals are initialized as producers ($p = 1$) that produce an amount of public good $E = 0.08$. Importantly, mutation probabilities of both p and E are set to zero. Instead, mutant individuals with a division of labor ($p = 0.5$) are introduced in the system by hand. The mutants have a probability of 0.5 to produce an amount of public good $\phi = 0.16$, and a probability of 0.5 to produce nothing. Even though the expected level of public-good production is the same for the individuals with and without a division of labor, the mean cost is approximately 25% lower for the individuals with a division of labor compared to the individuals without a division of labor (Equation (3) in Methods).

8.1.1 Software and code availability

The software used for the simulations was written in C. The code is available on Github:
https://github.com/ibouwman/altruism/tree/track-colonies for within-group experiments, and
https://github.com/ibouwman/altruism/tree/track-colonies-among for among-group experiments. Statistical analyses were performed in R version 4.1.1 and figures were made in Wolfram Mathematica.

8.2 Results

8.2.1 Competition within groups

Competition between lineages with and without a division of labor is simulated by introducing a cluster of 100 mutant individuals with a division of labor in the middle of a group of wildtype individuals without a division of labor. Notably, all individuals have the same expected level of public-good production. The frequency of individuals with a division of labor in the group was tracked over time. Mutants with a division of labor did not reach higher frequencies within a group than individuals without a division of labor under the same conditions (Welch two sample t-test, $P = 0.14$, $N = 30$).

8.2.2 Competition among groups

We consider two scenarios in which a fitter group can outcompete other groups. First, a fitter group can increase in size because the individuals in the group have a higher reproduction rate than the individuals in the neighboring groups. As a result, individuals in neighboring groups experience increased competition, so their reproduction rate declines. In turn, the individuals in the fitter group can reproduce even faster, because they experience less competition from the surrounding groups. Finally, one of the surrounding groups dies out and the fitter group splits. This scenario is simulated by initializing one group that fully consists of mutants with a division of labor in the habitat, while all other groups do not have a division of labor. Thus, there is no within-group variation. All groups have the same expected level of public-good production, but the mean cost of the group with division of labor is approximately 25% lower than the mean cost of the other groups. Next, we test whether the group with division of labor can outcompete the groups without a division of labor. In the experiments

performed here, the group with a division of labor never expands and hence does not outcompete the groups without a division of labor ($N = 20$).

In the second scenario, a group dies because it is taken over by cheaters and thus no longer has public good available ($E = 0$). Consequently, one of the six neighbors of the dying group can divide, and its offspring group fills the empty space left by the group that disappeared. We simulate this scenario by initializing one group with $E = 0$, while all other groups consist of individuals with $E = 0.08$. In order to test competition between groups with and without division of labor, three neighbors of the group without public good have a division of labor, while the other three groups do not have a division of labor. The scenario is tested for two different configurations: groups with and without division of labor alternate, or groups with a division of labor flock together (Figure S1). Next, we assess whether groups with a division of labor take over empty spots in the grid more often than groups without a division of labor. For both configurations considered here, groups with and without a division of labor take over empty spots equally often ($N = 10$).

8.3 Conclusions

The simulations presented above aim to provide different experimental setups to study within-group and among-group selection of division of labor in the two-dimensional model used in this study. Preliminary simulations with small sample sizes and a limited set of parameter values did not indicate a role for either within-group or among-group selection of division of labor. However, selection of a trait can be driven by small fitness differences. Therefore, simulations with larger sample sizes might reveal a fitness benefit of division of labor within or among groups.

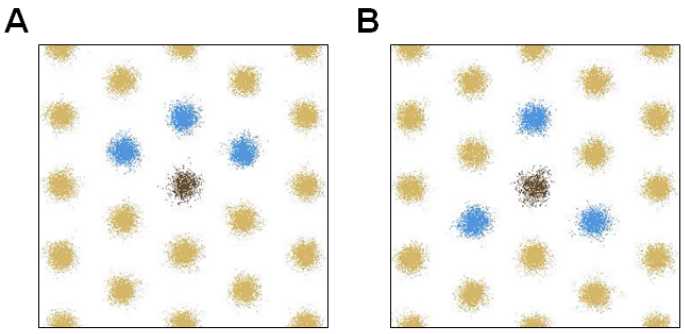

Figure S1: Experimental setup to test competition among groups. A group of defectors (brown, $E = 0$) is surrounded by three groups with a division of labor (blue, $p = 0.5$) and three groups without a division of labor (yellow, $p = 1$). Two different configurations are shown: groups with division of labor flock together (A), or groups with and without division of labor alternate (B). Except for the group of defectors, all groups in the field have the same expected production $E = 0.08$. There is no within-group variation.

9 Supplementary information

Video S1: Video showing the spatial positions of producers (black) and non-producers (yellow) over time, as depicted in Figure 3. Vertical axis ticks show the scales of altruism (left) and competition (right). The time interval between the shots of the video is 5000 generations. The video is available at:
https://drive.google.com/file/d/1YhcK2blj061IfldPmTcnF8isI3dFZgJa/view?usp=sharing

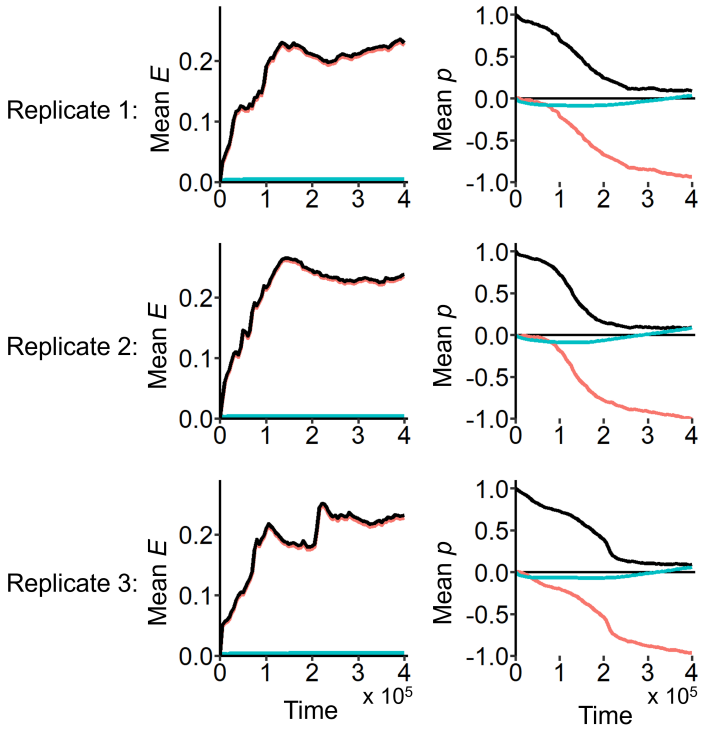

Figure S2: Replicates of simulations performed with the default parameters (Table 1). The mean expected production (E, left) and the mean probability to produce (p, right) are plotted over time, with the relative contributions of selection and drift (red) and transmission (blue).

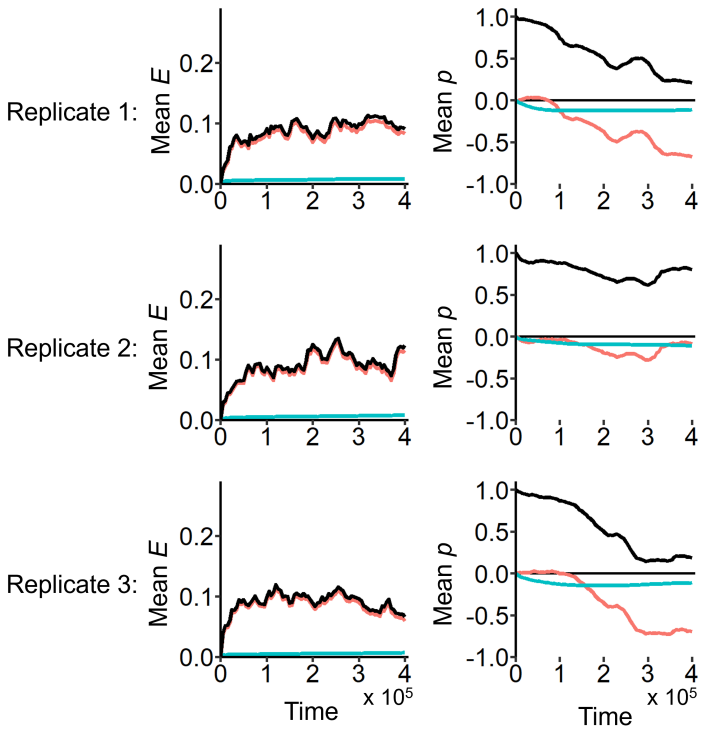

Figure S3: Replicates of simulations performed with $\alpha = 0.5$. The mean expected production (E, left) and the mean probability to produce (p, right) are plotted over time, with the relative contributions of selection and drift (red) and transmission (blue). Except for α, the default parameters in Table 1 were used.

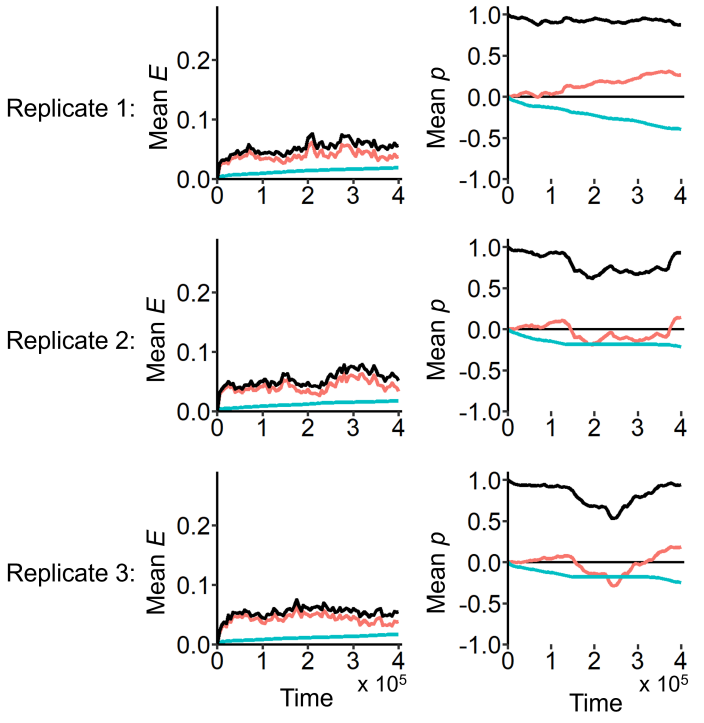

Figure S4: Replicates of simulations performed with a linear cost function ($\alpha = 1$). The mean expected production (E, left) and the mean probability to produce (p, right) are plotted over time, with the relative contributions of selection and drift (red) and transmission (blue). Except for α, the default parameters in Table 1 were used.

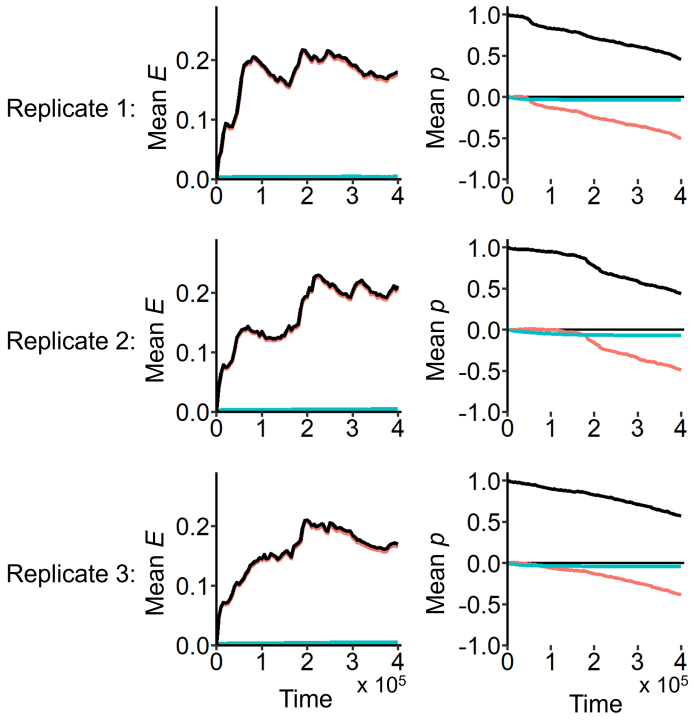

Figure S5: Replicates of simulations performed with the default cost function ($\alpha = 0.1$), but a lower mutation probability in p ($\mu_p = 0.001$). The mean expected production (E, left) and the mean probability to produce (p, right) are plotted over time, with the relative contributions of selection and drift (red) and transmission (blue). Except for μ_p, the default parameters in Table 1 were used.

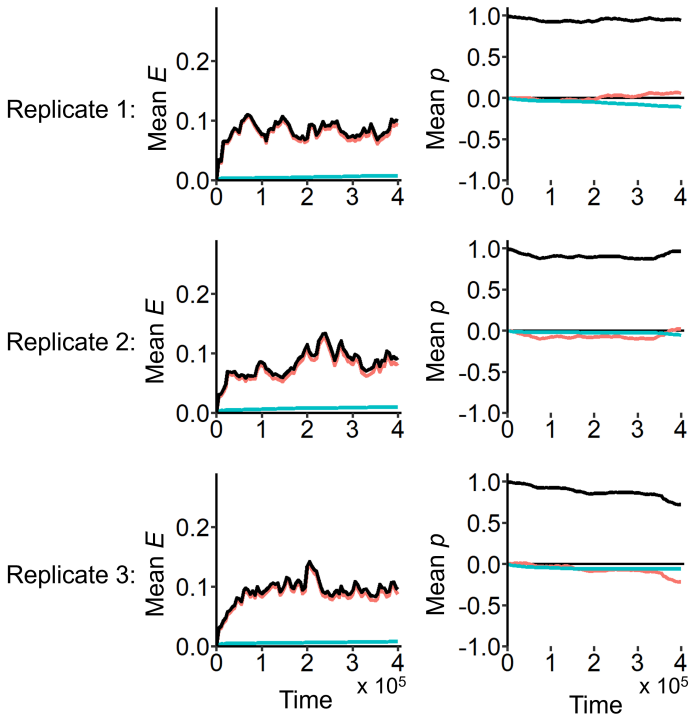

Figure S6: Replicates of simulations performed with $\alpha = 0.5$ and a lower mutation probability in p $\mu_p = 0.001$. The mean expected production (E, left) and the mean probability to produce (p, right) are plotted over time, with the relative contributions of selection and drift (red) and transmission (blue). Except for α and μ_p, the default parameters in Table 1 were used.

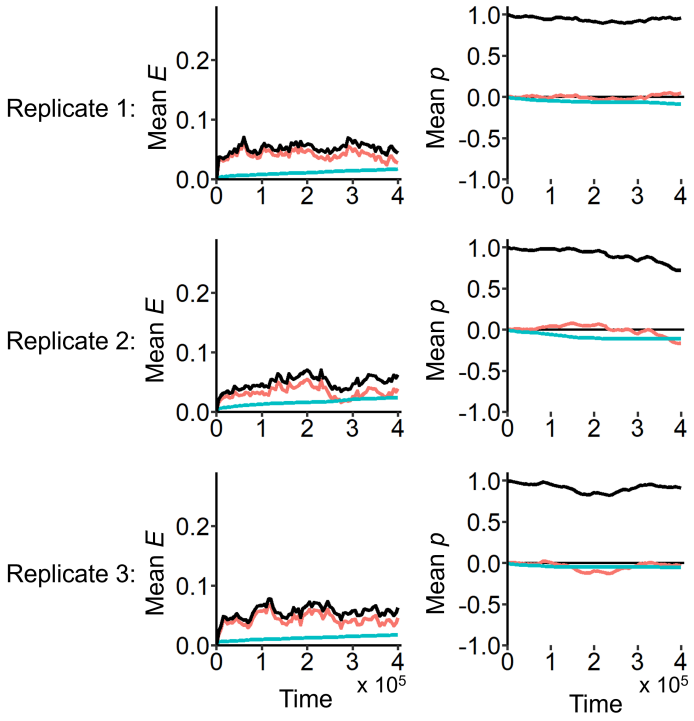

Figure S7: Replicates of simulations performed with a linear cost function ($\alpha = 1$) and a lower mutation probability in p ($\mu_p = 0.001$). The mean expected production (E, left) and the mean probability to produce (p, right) are plotted over time, with the relative contributions of selection and drift (red) and transmission (blue). Except for α and μ_p, the default parameters in Table 1 were used.